CLICK-EVOKED OTOACOUSTIC EMISSIONS FOR THE ASSESSMENT OF AUDITORY FILTER TUNING AT SUPRA-THRESHOLD LEVELS

Sarah Verhulst[1], and Stefan Raufer[1,2]

[1]Cluster of Excellence "Hearing4all", Dept. of Medical Physics and Acoustics, Carl von Ossietzky University Oldenburg, DE-26129 Oldenburg, Germany

[2]Program in Speech and Hearing Bioscience and Technology, Harvard University, Cambridge, MA 01238, USA
e-mail: sarah.verhulst@uni-oldenburg.de

Reliable estimates of supra-threshold (60-80 dB SPL) filter tuning are necessary to understand auditory processing of speech. However, existing approaches suffer from methodological limitations that require high suppressor tone levels (psychoacoustics) or an assumption of linearity (otoacoustic emission group delay) to estimate human auditory filter tuning at higher stimulus levels. We propose a method based on temporal suppression (or, forward masking) of click-evoked otoacoustic emissions (CEOAE) to assess human supra-threshold filter tuning, with the underlying thought that the basilar-membrane impulse-response duration can be estimated from the spectral peaks in an individual's CEOAE spectrum and how they interact with the forward masker. We derived tuning values for 11 subjects, recorded at 65 dB peSPL, and observed a clear frequency-dependence of tuning that resembles that of near-threshold otoacoustic emission tuning estimates. The mean Q values were 8.4 at 1 kHz, and sharpened to 12.1 at 4 kHz. We compared our method to psychoacoustic tuning curves (PTC) measured in the same subjects at equivalent levels, and found that PTC tuning was generally sharper than that obtained using the OAE methods.

1. Introduction

Click-evoked otoacoustic emissions are elicited by a broadband stimulus and thus provide information about cochlear processing at several cochlear locations at once. Coherent reflection filtering theory links the frequencies in the CEOAE spectra to their corresponding cochlear locations [1], allowing the study of how these spectral components change as a function of stimulus alterations. Several studies have investigated temporal suppression of CEOAEs by placing a suppressor click 0-10 ms before a test click that elicited a CEOAE, and found that the root-mean-square (rms) level of the CEOAE is significantly reduced for inter-click intervals (ICIs) between 0-8 ms as a consequence of the suppressor [2,3,4,5].

If temporal CEOAE suppression is a consequence of how the spectral peaks in the CEOAE spectrum are suppressed by a specific ICI, it could be possible to estimate the basilar-membrane (BM) impulse response duration from studying the interactions between a spectral CEOAE component and the ICI. The underlying thought is that a suppressor at a specific ICI would only be able to suppress a spectral component in the CEOAE for as long as the corresponding impulse responses of both clicks interact. The ICI for which the interaction is gone (i.e., decay time or release of suppression) would then inform about the BM impulse response duration. Using filter theory, the bandwidth of the corresponding filter can be estimated using $BW = \frac{1}{Decay\ time[ms]}$ and the associated filter tuning (Q) is found by dividing the frequency of the CEOAE spectrum (i.e., characteristic frequency; CF) by that bandwidth: Q=CF/BW.

The proposed method benefits from that no assumptions of cochlear linearity need to be made to derive the tuning estimates from impulse response durations (that may reflect nonlinear compressive auditory filtering). Secondly, because the adopted method to extract tuning relies on nonlinear interactions between CEOAEs to temporally spaced clicks, it is suited to study the supra-threshold level range of human auditory filter tuning. The method could thus form a valid complementary method to existing tuning estimation methods at low stimulus levels (stimulus-frequency OAE e.g. [6], or psychoacoustic masking masked tuning e.g. [7]).

2. Materials and Methods

A total of 11 adults (Ages: 19-29, mean: 23, 2 male, 9 female) participated in the experiments. Hearing status was tested and found normal (<20 dB HL) at audiometric frequencies between 0.125-8 kHz. Participants were provided with written and oral information about the experiment and were paid for their participation. Experiments were approved by the ethics commission of Oldenburg University (Kommission fuer Forschungsfolgenabschaetzung und Ethik).

2.1 CEOAE Recordings

CEOAE recordings were conducted in a double-walled sound-attenuating booth at the Acoustics Lab of Oldenburg University where participants were seated in a comfortable chair and instructed to move as little as possible. Stimuli were played back over ER-2 in-ear loudspeakers that were driven by a Tucker-Davis Technologies (TDT) HB7 headphone driver attached to a RME Fireface UC external sound card. OAEs were recorded using an Etymotic Research (ER), Inc. ER-10B+ Low-Noise Mic System and via the sound card stored in digital format. Up to 41 ICI conditions were measured within one 75-90 min session. In the very first session, and to check whether the subject had CEOAEs, conditions between 0 and 8 ms ICI were recorded in steps of 0.5 ms at a level of 65 dB peSPL. In the following sessions, ICI were measured in steps of 0.1 ms. ICI conditions were randomized across subjects and sessions.

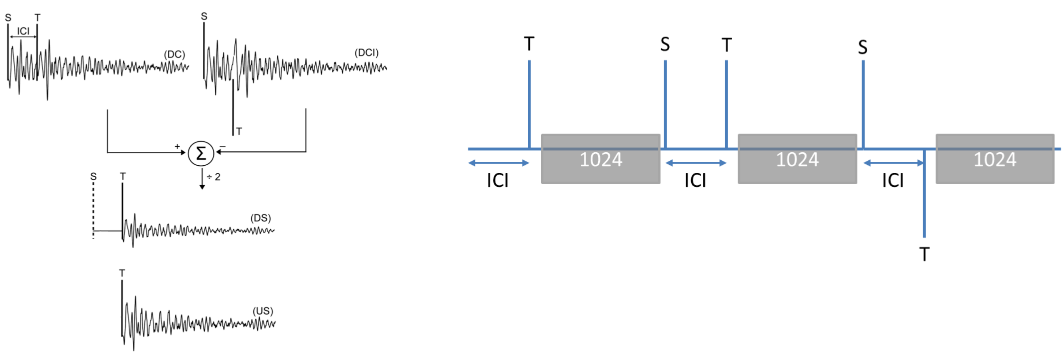


Figure 1. Left: CEOAE recording method of temporal suppression. A derived suppressed CEOAE waveform (DS) was obtained by subtracting a double-click inverted (DCI) condition from a double-click condition (DC) and by dividing the resulting waveform by 2. The DS waveform contained nonlinear influences of the ICI on the CEOAE waveform while having removed the linear interactions of the suppressor click (S) on the response to the test-click (T). **Right:** Schematic of one out of 1200 repeated stimulus blocks, containing the test-click alone, and a suppressor-click preceding the test-click with a certain ICI. The polarity of the test-click is altered in each double-click condition. The recording and analysis window for each test-click was fixed to 28 ms.

Calibration was performed by inserting the ER-2 probe in a 2-cc cavity attached to a B&K type 2669 ear simulator attached to a B&K type 2610 analog sound level meter. The clicks lasted 80 µs and Fig.1 illustrates how temporal suppression yields a derived-suppressed (DS) waveform of which the spectrum can be compared to the unsuppressed (US) spectrum to study temporal suppression of specific CEOAE components. Spectra were calculated for a 21 ms window that started 5-6 ms after the onset of the test-click, depending on subject specific decay times of the linear portion of the OAE.

2.2 Psychoacoustic Tuning Curves

Psychoacoustic tuning curves (PTCs) were obtained in each listener for up to 4 frequencies, each corresponding to a spectral peak in the CEOAE spectrum using the same sound delivery system as for the OAE experiments. A modified forward masking method based on [8] was used in which the 50 ms probe tone level was kept fixed, and a pure-tone masker of 500 ms length that ended 10 ms before the probe tone was varied in level to determine the masker level necessary to just detect the probe tone using an adaptive two alternative forced choice (AFC) tracking procedure [9]. Each condition was repeated 3 times, and 8 reversals determined the threshold at the lowest step size of 1 dB. The tested masker frequency conditions were randomized throughout the 90 min recording session, and subjects were provided with visual feedback of the correctness of their answers.

To construct a PTC, masker frequencies ranged between 0.85 and 1.07 times the probe tone center frequency. We measured 3 masker frequency conditions above and below the tone frequency, and one on-frequency condition. The exact frequencies of the maskers were set to obtain a dynamic range of the masker level of at least 10 dB measured from the tip of the tuning curve (on-frequency condition) to lower and higher masker frequencies. The probe-tone level was chosen to reflect as closely as possible the amount of energy the 65 dB peSPL click would produce inside a given auditory filter using a bandwidth compensated level procedure [10]:

$$(1) \qquad L_{bwc} = L_{pk} - 20\log(1+r),$$

where L_{pk} denotes the peSPL level of the click (65 dB), and r=BW\DF. DF was set to 74 Hz at a 1.2 kHz center frequency [9] and was extrapolated to other frequencies using:

$$(2) \qquad DF = 74Hz * (f_c/1.2kHz)\verb|^|0.7,$$

with the exponent following from the Q_{ERB} exponent for human tuning as a function of center frequency [6].

3. Results and Discussion

3.1 Temporal Suppression of CEOAE components

Figure 2 shows unsuppressed (US) CEOAE spectra (red) for one of the tested listeners plotted along with derived suppressed (DS) CEOAE spectra (black) for different ICIs. Whereas reduced spectral peaks are clearly observed from the spectral peaks in the 0 and 1 ms ICI condition, the larger ICI conditions only partly suppress the frequency components of the CEOAE spectrum. More specifically, shorter ICIs were able to suppress all frequency components, whereas longer ICIs only suppressed lower frequency spectral components. This observation was made for all listeners in this study suggesting a suppression relationship between the ICI and frequency of the CEOAE component. This was quantified further by analysing the temporal suppression amount for all identified spectral CEOAE peaks in all listeners for 0.1 ms ICI steps. Spontaneous otoacoustic emissions were excluded from the analysis. An example of temporal suppression patterns for two CEOAE components is shown in Figure 3, and shows that suppression patterns for the high frequency components decays down for shorter ICIs than for the lower frequency component. For each of the CEOAE components, we observed that temporal suppression increased for the first ICIs after which a maxi-

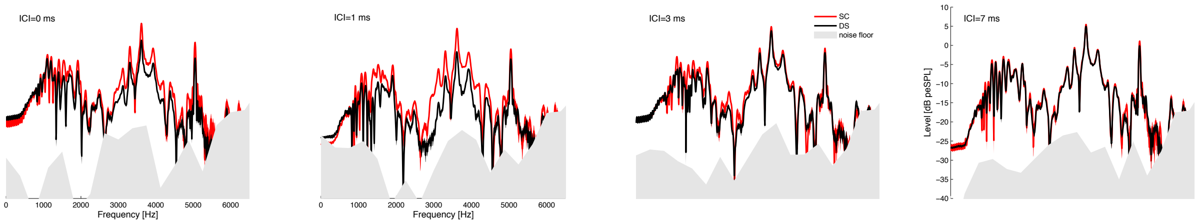


Figure 2. CEOAE spectra for one representative subject for ICIs beween 0 and 7 ms. The red line represents the spectrum of the unsuppressed (US) CEOAE, whereas the black line is that of the derived suppressed (DS) CEOAE. The gray area in each panel is the estimated noise floor of each condition. For low ICIs, most spectral components are suppressed and as ICI increases, high-frequency components are released from suppression before low frequencies are.

mum was reached and a decaying trend was observed. The small oscillations observed in the temporal suppression could be due to the phase relationship between the ICI duration and the period of the underlying BM impulse response. If ICIs are only able to suppress CEOAE components for as long as the underlying BM impulse responses to the double click stimuli are able to interact, the release of suppression (i.e., the ICI for which temporal suppression disappears) should inform about their duration. The release of suppression (i.e., decay time) was established from temporal suppression patterns of identified spectral CEOAE components by fitting an exponentially decaying least-square fit to the patterns, starting from the ICI of maximal suppression toward higher ICIs (c.f., Fig. 3). The decay time for temporal suppression of each spectral CEOAE component corresponded to three time constants of this exponential fit, i.e. 95% change in state from the maximum.

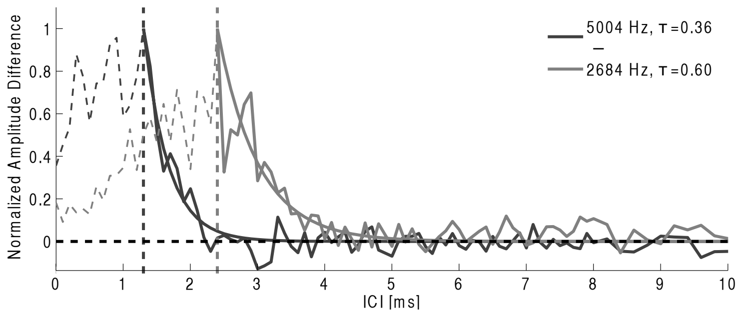


Figure 3. Normalized suppression as a function of ICI for two CEOAE frequencies of one individual as a function of ICI. The solid lines indicate an exponentially decaying Least-Square fit to the data, starting from the ICI where suppression exceeds its maximum (dashed vertical lines). Time constants (τ) of the exponential functions are given in the legend. The time-constant of suppression is smaller for higher CEOAE components consistent with their shorter associated basilar-membrane impulse response duration.

3.2 CEOAE-derived Filter Tuning

If the decay time of temporal suppression patterns reflect the underlying BM impulse response duration, the associated BM filter bandwidth (BW) and tuning (Q) can be derived using $BW = \frac{1}{Decay\ time[ms]}$ and Q=CF/BW. Figure 4 shows the derived tuning values for all spectral CEOAE components considered in this study. In correspondence to studies that investigated human auditory filter tuning at low stimulus levels as a function of characteristic frequency [6,7,11], we fitted a po-

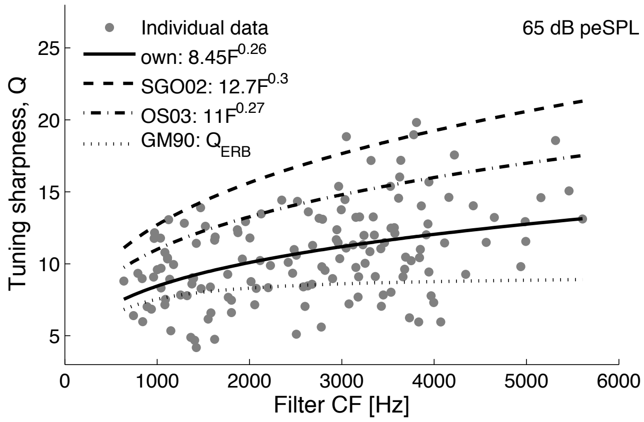


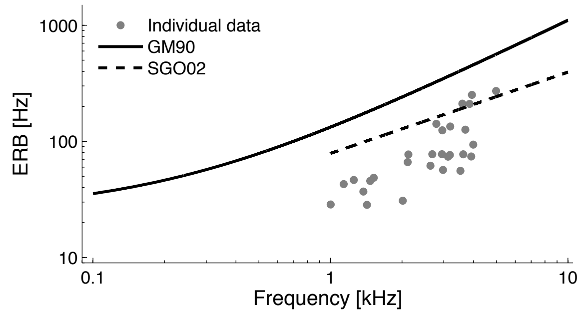
Figure 4. Estimated tuning sharpness as a Function of CF for levels of 65 dB peSPL. Least-Square power-law fits are provided. For comparison, the Q_{ERB} in [GM90; 11] (dotted line), and the sharpness of tuning estimates of [SGO02; 6] (dashed line), and [OS03; 7] (dash-dot line) are shown. The exponent of our power-law fit is in range of those found in other studies, even though our tuning values are generally lower than those in the [6,7] studies and sharper than those reported in [11].

power law function $Q(f) = 8.45f[kHz]^{0.26}$ to our data and found that the exponent of this function closely matches the 0.27-0.3 exponents found in other studies [6,7]. This finding supports the hypothesis of this study that assumed that the release of temporal suppression for spectral CEOAE components is related to the underlying BM impulse response duration.

However, the exact values of tuning for any specific frequency were found to be different than those reported in other studies. At 1 kHz, we report a value of 8.45 that is less sharp than that derived from low-level stimulus-frequency OAE group delays [12.7; 6], and that derived from psychoacoustic forward masking experiments [11; 7]. This can be partly explained by the difference in tested stimulus levels. Whereas we used supra-threshold levels (65 dB peSPL), the OAE and psychoacoustic methods estimated tuning at a stimulus level of 40 dB, where the filters are expected to be sharper and associated impulse responses longer. Additionally, because our method estimates the BM impulse-response duration from the nonlinearly overlapping part of BM impulse responses (see method in Fig.1), it is possible that our method underestimates the actual BM impulse response duration slightly. In contrast, our method reports sharper tuning values than those reported in [11], because the latter study relies on a simultaneous masking method that generally yields less sharp tuning values than its forward masking psychoacoustic alternatives [12]. Lastly, the exact decay time criterium adopted in the filter bandwidth estimated may shift the tuning curves more or less vertically, while this is not expected to not influence the variation of tuning across frequency.

3.3 Psychophysical Tuning Curves

To test how our proposed method of supra-threshold tuning relates to established psychoacoustic methods for estimating auditory filter tuning, we adopted a forward masking method for probe tones that matched closely the spectral level of the evoking click used in the CEOAE method (Eqs. 1, 2). Psychoacoustic tuning curves were obtained from up to four spectral CEOAE frequencies per listener at probe tone levels corresponding to the energy of the 65 dB peSPL click in the same area (see Eq. (1)). Results are shown in the right panel of Figure 5. The tuning values were obtained from determining the equivalent-rectangular bandwidth (ERB) from the PTCs and by applying Q_{ERB}=CF/ERB. As observed from the left panel of Fig. 5, the psychoacoustic derived ERBs are much narrower than observed in both forward and simultaneously masked studies [7, 11]. This ob-



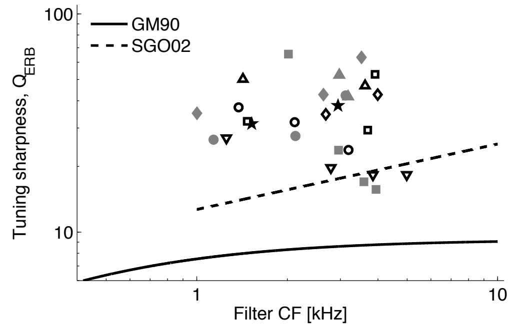


Figure 5. Left: Equivalent rectangular bandwidths (ERBs) derived from the PTCs for center frequencies ranging from 1-5 kHz. Shown are 28 frequencies from 10 subjects (2-4 per subject). For comparison, data obtained from GM90 [11] and SGO02 [6] are shown. **Right:** Estimated Q_{ERB} values from the same data as in the left panel. Each marker (color) represents one individual.

servation is surprising, especially because it yields tuning values that are much sharper than observed in even the sharpest reported tuning values of other studies [6]. The stimulus configuration, i.e. the pure tone forward masking paradigm, is a possible contributor to the observed sharp tuning values. Alternatively, it may be that hearing sensitivity near spectral CEOAE peaks is slightly better than that of neighbouring frequencies in line with how hearing sensitivity is better near SOAE frequencies [13].

Even if these suggestions would explain the difference in tuning across the PTC and CEOAE method, the methods are fundamentally different. The psychophysical method estimates a tuning curve for sustained pure-tones by making use of a fixed 10 ms forward masker, whereas the CEOAE method estimates tuning from BM impulse response durations, i.e. a estimation method to transient stimuli. A good correspondence across these methods is thus not expected a priori.

4. Conclusion

The proposed CEOAE method to assess supra-threshold human auditory filter tuning yielded tuning values across CF in close agreement with power-law exponents found in other human SFOAE [6] and psychoacoustic [6,7] tuning studies. This supports the relation between our measure of temporal suppression and duration of the BM impulse responses associated with the peaks of the CEOAE spectrum. The absolute values of tuning for a specific CF were found to lie between those found in psychoacoustic forward [6,7] and simultaneous masking [11] studies, and would need a precise estimate of the spectral energy that was available at a specific BM location to allow more detailed comparisons to methods that adopted pure-tones as stimuli to assess tuning. Our method benefits from that it is applicable at higher stimulus levels where the system behaves nonlinear, and therefore offers a window to studying human auditory filter tuning at supra-threshold levels.

Acknowledgement

The work was supported by DFG Cluster of Excellence EXC 1077/1 “Hearing4all”. The authors would like to thank Sandra Tolnai for interesting discussions related to this work.

REFERENCES

1 Shera, C.A., and Guinan, J.J.Jr. Evoked otoacoustic emissions arise by two fundamentally different mechanisms: A taxonomy for mammalian OAEs, *Journal for the Acoustical Society of America*, **105** (2), 782- 798, (1999).

2 Kemp, D.T. and Chum. Properties of the generator of stimulated acoustic emissions, *Hearing Research,* **2** (3), 213-232, (1980).

3 Kapadia, S., and Lutman, M.E. Nonlinear temporal interactions in click-evoked otoacoustic emissions. I. Assumed model and polarity-symmetry, *Hearing Research*, **146**, 89-100, (2000a).

4 Kapadia, S., and Lutman, M.E. Nonlinear temporal interactions in click-evoked otoacoustic emissions. II. Experimental data, *Hearing Research*, **146**, 101-120, (2000b).

5 Verhulst, S., Harte, J.M., and Dau, T. Temporal suppression of the click-evoked otoacoustic emission level-curve, *The Journal of the Acoustical Society of America*, **129** (3), 1452-1463, (2011).

6 Shera, C.A., Guinan, J.J.Jr., and Oxenham, A.J. Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements, *PNAS*, **99** (5), 3318-3323, (2002).

7 Oxenham, AJ., and Shera, C.A. Estimates of Human Cochlear Tuning at Low Levels Using Forward and Simultaneous Masking, *Journal of the Association for Research in Otolatyngology*, **4**, 541-554, (2003).

8 Moore, B.C.J. Psychophysical tuning curves measured in simultaneous and forward masking, *Journal of the Acoustical Society of America*, **63** (2), 524-532, (1978).

9 Levitt, H. Transformed Up-Down Methods in Psychoacoustics, *The Journal of the Acoustical Society of America*, **49**, 467-477, (1971).

10 Kalluri, R., and Shera, C.A. Near equivalence of human click-evoked and stimulus-frequency otoacoustic emissions, *The Journal of the Acoustical Society of America*, **121** (4), 2097-2110, (2007).

11 Glasberg, B.R., and Moore, B.C.J. Derivation of auditory filter shapes from noteched noise data, *Hearing Research*, 47, 103-138, (1990).

12 Eustaquio-Martin, A., and Lopez-Poveda, E. A. Isoresponse and Isoinput Estimates of Cochlear Filter Tuning, *Journal of the Association for Research in Otology*, **12**, 281-299, (2011).

13 Long, G. R., and Tubis, A. Investigations into the nature of the association between threshold microstructure and otoacoustic emissions, *Hearing Research*, **36**, 125–138 (1988).